Interplay between stochasticity and negative feedback leads to pulsed dynamics and distinct gene activity patterns

Interplay between stochasticity and negative feedback leads to pulsed dynamics and distinct gene activity patterns

Samuel Zambrano*
San Raffaele University, Via Olgettina 58, 20132 Milan, Italy
and Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milan, Italy

Marco E. Bianchi
San Raffaele University, Via Olgettina 58, 20132 Milan, Italy

Alessandra Agresti
Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milan, Italy

Nacho Molina†
SynthSys Centre, University of Edinburgh, Mayfield Road, EH9 3JD Edinburgh, United Kingdom
(Received 14 November 2014; revised manuscript received 13 March 2015; published 14 August 2015)

Gene expression is an inherently stochastic process that depends on the structure of the biochemical regulatory network in which the gene is embedded. Here we study the dynamical consequences of the interplay between stochastic gene switching and the widespread negative feedback regulatory loop in a simple model of a biochemical regulatory network. Using a simplified hybrid simulation approach, in which only the gene activation is modeled stochastically, we find that stochasticity in gene switching by itself can induce pulses in the system, providing also analytical insights into their origin. Furthermore, we find that this simple network is able to reproduce both exponential and peaked distributions of gene active and inactive times similar to those that have been observed experimentally. This simplified hybrid simulation approach also allows us to link these patterns to the dynamics of the system for each gene state.

DOI: 10.1103/PhysRevE.92.022711 PACS number(s): 87.18.Cf, 87.16.Yc, 87.18.Tt, 87.18.Vf

I. INTRODUCTION

Cells need to provide an adapted response to external stimuli, which requires the production of adequate proteins following different temporal patterns. This is achieved through biochemical networks in which a stimulus triggers a cascade of reactions that eventually lead to the activation of transcription factors, proteins that activate or repress the expression of specific gene sets. Thus, the temporal regulation of gene activity will be determined by the structure of the network in which the gene is embedded [1]. A common regulatory structure is the negative feedback loop, in which a transcription factor activates the production of a protein that contributes to its own inhibition. This motif regulates the activity of important transcription factors such as NF-κB [2] and p53 [3] and has been shown to give rise to pulses in the concentration of the proteins of the network (see, e.g., [4,5]) as predicted by mathematical models [6]. The role of this pulsed dynamics is not fully understood though suggestions include stability and reliability in protein production [7] and a role in determining the cell fate [8].

Gene expression is an intrinsically noisy process [9] and a simple model in which the gene state switches randomly between active and inactive states has been proposed [10]. Such a model is able to fit experimental data of steady-state distributions of gene expression levels in single cells [11]. For the gene activity dynamics, this simple model predicts exponential distributions of the gene active and inactive times. However, recent experiments *in vivo* suggest that some genes show instead peaked distributions [12–14]. These observations can be reproduced using multistate gene models mirroring the multiple steps of gene activation [15], but these distributions could also arise from the interplay between the stochastic gene activity and the structure of the regulatory network in which the gene is embedded. Important analytical insights have been gained into how such interplay shapes the steady-state distribution of gene expression levels under quite general contexts [16,17], but less is known about the role of this interplay in the dynamics. Several insights have been gained recently by showing the emergence of oscillations when a gene is an autorepressor [18] and of noise-enhanced persistence of biochemical species [19] and how stochasticity dephases genetic oscillators [20]. A major obstacle in this context is the difficulty of treating analytically the dynamics of the nonlinear stochastic systems involved (often involving species with very low copy numbers). For this reason, we are far from having a complete picture of the type of regulation that emerges from such interaction.

In this paper we provide further insights into this interplay by describing the dynamics emerging in a simple network with a stochastic gene switching and the common negative feedback loop. In particular, we describe the dynamics and the gene activity patterns that this kind of network produces by using a simplified hybrid simulation approach, in which only the gene switching is simulated as a stochastic process while the remaining variables of the network are modeled

*zambrano.samuel@hsr.it
†nacho.molina@ed.ac.uk

through ordinary differential equations. Our simplified hybrid simulation approach shows that stochastic gene switching is responsible for most of the dynamical variability of the system and that it leads to the emergence of pulses in the network, although the deterministic simulation predicts a steady state. Moreover, we show that even for our simple biochemical network, distinct dynamical patterns of gene activity can arise, reminiscent of those observed in a number of experiments and how the simplified hybrid simulation approach allows us to gain analytical insight into their origin. We discuss the implications of our results in Sec. VI.

II. SIMPLE GENE CIRCUIT WITH NEGATIVE FEEDBACK

The model considered here is shown in Fig. 1(a). This model adds a layer of regulation to the one proposed in Ref. [7] and is a simplified version of the biochemical network of NF-κB [21]. With this simple model we do not intend to provide a detailed description of NF-κB oscillations, but instead gain an understanding of the role of the interaction between stochastic gene switching and the presence of negative feedback.

Our model consists of a gene that can be active G or inactive $\bar{G}$ and an activator A that, similarly to NF-κB [21], can activate the gene:

$$\bar{G} + A \xrightarrow{k_{\text{on}}} G + A. \tag{1}$$

When the gene is active, the inhibitor protein I is produced and we summarize transcription and translation as

$$G \xrightarrow{k_I} G + I. \tag{2}$$

As in the NF-κB biochemical network [21], this inhibitor provides the negative feedback by both contributing to the gene's inactivation

$$G + I \xrightarrow{k_{\text{off}}} \bar{G} + I \tag{3}$$

and forming a complex with A, $A{:}I$, that cannot activate the gene any longer, a complex that can also dissociate

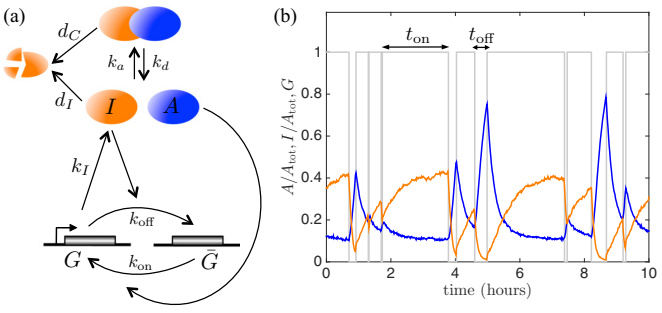

FIG. 1. (Color online) Dynamics of a network with stochastic gene switching and a negative feedback loop. (a) Model diagram showing the activator A, the inhibitor I, and the gene that can be active G or inactive $\bar{G}$. (b) Stochastic simulation of a trajectory of the system showing the free activator fraction A/A_{tot} [blue (gray)], the gene state G (step gray line), and the relative amount of inhibitor I/A_{tot} [orange (light gray)] displaying pulses. For the gene state, an active time t_{on} and an inactive time t_{off} are displayed. Simulations were performed with the parameters $d_C = 1.5 \times 10^{-4}$ s^{-1}, $k_a = 5.1 \times 10^{-6}$ s^{-1}, $k_d = 1.4 \times 10^{-3}$ s^{-1}, $d_I = 2.8 \times 10^{-4}$ s^{-1}, $k_I = 11.2$ s^{-1}, $k_{\text{off}} = 1.4 \times 10^{-7}$ s^{-1}, and $k_{\text{on}} = 5.8 \times 10^{-7}$ s^{-1}.

spontaneously. Thus,

$$A{:}I \underset{k_d}{\overset{k_a}{\rightleftharpoons}} A + I. \tag{4}$$

Finally we impose that the inhibitor undergoes degradation both in the free form

$$I \xrightarrow{d_I} \emptyset \tag{5}$$

and when it forms a complex with the activator, so

$$A{:}I \xrightarrow{d_C} A. \tag{6}$$

In what follows we use the same letters both for the names of the biochemical species and for their copy numbers. For the sake of simplicity we consider that we have only one copy of the gene, so $G + \bar{G} = 1$, and that the total amount of A (free and bound to I) is constant and equal to A_{tot}, as for NF-κB [21].

In Fig. 1(b) we show a stochastic trajectory of the system obtained using the Gillespie algorithm [22]. It exhibits pulses of the free activator A and the rest of variables, as observed in different networks with negative feedback [6]. Due to the simplifications behind the present model, it is not possible to provide an estimation to some parameters (e.g., parameter k_I summarizes the many steps of the transcription and translation process). Thus, the parameters used were obtained from the slightly more complex deterministic model of Ref. [21] and adjusted to obtain $O(10^4)$ copies of each protein and inhibitor half-life and activator intrapeak timing of the order of 1 h, the typical time scale of these biological oscillators [6]. Pulses are spiky as in certain models of oscillations of NF-κB [6], but it will soon become clear that this is due to the simplicity of our network. Furthermore, as we will show in the following sections, the dynamics of the network are qualitatively analogous independently of the parameters considered, while distinct gene activity patterns arise when varying the parameters up to two orders of magnitude.

III. SIMPLIFIED HYBRID SIMULATION

Stochastic simulations for large models containing nonlinearities can be computationally expensive. Under certain assumptions of a high number of molecules and low noise (see the Appendix), a Langevin equation [23], a more tractable approach, can be used to approximate the stochastic dynamics, although exact analytic results can only be obtained for linear biochemical networks [24]. In the limit of a very large number of molecules, when fluctuations can be neglected, ordinary differential equations derived from mass-action kinetics are traditionally used to describe the deterministic dynamics of the system.

Both approaches are inadequate when species with low copy numbers are present (in our case, G). To overcome this problem, the so-called hybrid simulations, in which part of the reactions are modeled by a Langevin equation [25] while the rest are modeled as discrete stochastic processes, have been proposed [26]. The exact criteria to decide which modeling approach is appropriate for each reaction are provided in the Appendix. Essentially, a biochemical reaction can only be modeled using the Langevin equation if it is likely to occur

in short time intervals [Eq. (A1)] and if it involves relatively small changes in the copy numbers of the biochemical species involved [Eq. (A2)]; in short, only if the considered species can be approximated by continuous variables. The resulting hybrid simulations (combining Langevin equations and discrete stochastic processes) faithfully mimic the fully stochastic simulations while significantly reducing the computation time.

In this context, a *simplified hybrid* simulation approach has been becoming increasingly popular in order to simulate complex models of cell signaling and gene expression [27–30]. In practice, this simplified scheme essentially models as ordinary differential equations the processes that should be modeled as Langevin equations using a hybrid scheme (see [26] and our Appendix) and the rest as discrete stochastic processes. However, very little attention has been paid to the accuracy of this kind of simulation. For this reason, we study here the simple network proposed in Sec. II using this simplified hybrid simulation approach, primarily to assess its validity. More importantly, in this scheme only the gene activity is simulated as a discrete stochastic process, so this approach will allow us to to isolate and identify in a precise way the role of stochastic gene activity in the dynamics. Furthermore, using this approach we will be able to use ideas from dynamical systems to gain insight into the dynamics of the system.

Hence, considering the reactions described in Sec. II, the only variables whose evolution could be approximated by a continuous Langevin equation in the hybrid scheme are A and I (those with high copy number). Hence using a simplified hybrid simulation approach, we describe their evolution using the equations

$$\frac{dA}{dt} = -k_a A I + (k_d + d_C)(A_{\text{tot}} - A), \tag{7}$$

$$\frac{dI}{dt} = -k_a A I + k_d (A_{\text{tot}} - A) - d_I I + k_I G. \tag{8}$$

This nonlinear dynamical system is driven by the stochastic process of gene switching, which cannot be approximated faithfully by a continuous variable, and given by reactions (1) and (3), which we can summarize now as

$$\bar{G} \underset{k_{\text{off}} I(t)}{\overset{k_{\text{on}} A(t)}{\rightleftarrows}} G, \tag{9}$$

so the switching rates will depend on time through the continuous variables $A(t)$ and $I(t)$. As we will see below, the point of view provided by this scheme can help us gain an understanding of the system's dynamics.

For the gene state G we can write down the chemical master equation

$$\frac{dP(G)}{dt} = k_{\text{on}} A P(\bar{G}) - k_{\text{off}} I P(G). \tag{10}$$

A simplified hybrid numerical simulation is performed using a deterministic integrator for Eqs. (7) and (8) and switching the value of G between $G = 0$ and 1 following the Gillespie algorithm, as prescribed in Ref. [30]. For a fully deterministic simulation of the model using mass-action kinetics it is enough to add to Eqs. (7) and (8) the equation

$$\frac{dG}{dt} = k_{\text{on}} A \bar{G} - k_{\text{off}} I G, \tag{11}$$

which in the equilibrium ($\frac{dG}{dt} \approx 0$) leads to a Michaelis-Menten-like equation for G [1]. In what follows we show how the simplified hybrid simulation provides an insightful point of view to characterize the dynamics arising in our simple biochemical network.

IV. ROLE OF STOCHASTIC GENE ACTIVITY IN THE DYNAMICS

In order to understand to what extent the simplified hybrid simulation approach is able to capture the dynamics of the network, in Fig. 2(a) we show the evolution in time of the free activator A obtained for fully stochastic, simplified hybrid, and deterministic simulations. We can observe that the pulses obtained for the fully stochastic simulations and the simplified hybrid simulations are very similar. Interestingly, we observe that the deterministic simulations obtained by integrating Eqs. (7), (8), and (11) lead to the convergence of the system to a steady state. From this we conclude that stochastic gene switching by itself can induce pulses in the network. Of course, this does not imply that pulsed dynamics are always due to the stochasticity in gene activation. However, our simple model highlights the importance of such stochasticity in the global dynamics arising in these systems. This is another example of how stochasticity can induce pulses in contexts where deterministic models predict steady states, as observed in models of population dynamics [31] and excitable systems [32]. Contrary to those examples, here we are not dealing with a Langevin equation in which white noise can induce oscillations: It is precisely the fact that gene switching cannot be approximated as a reaction in a Langevin equations that makes pulses arise.

Our simulations also show that stochastic gene activity is responsible for most of the variability of the system. In Figs. 2(b) and 2(c) we show the distributions of A and I, represented by their probability density ρ, for stochastic and simplified hybrid simulations, which are nearly indistinguishable. On the other hand, by using a simple peak detection algorithm we can detect the timing between two consecutive peaks T and their amplitude A_{peak}. For these calculations we consider only peaks of at least 10% of A_{tot}, the order of magnitude that can be detected in experiments of activator dynamics such as NF-κB [5]. The distributions of these magnitudes are shown in Figs. 2(d) and 2(e), respectively, and are again nearly indistinguishable: This confirms the crucial role of stochastic gene activity in the dynamic variability of the system and the ability of simplified hybrid simulation to mimic the fully stochastic simulations (in drastically shorter computation times).

To obtain a wider perspective on how accurately the simplified hybrid simulations can mimic the fully stochastic simulations generated by the Gillespie algorithm, we performed simplified hybrid and stochastic simulations of the dynamics of our simple model, varying the parameters one order of magnitude above and below the parameters used above, and compared the difference in the distributions of A, I, A_{peak}, and T. To do this, we binned each magnitude to obtain discretized probability distributions from simplified hybrid and stochastic simulated traces of 2000 h [$P_H(i)$ and $P_S(i)$], adding a pseudocount $\alpha = 10^{-6}$ to avoid zeros in the discrete probabilities. Finally, we used the Kullback-Leibler divergence D_{KL} [33] to quantify how much information is lost

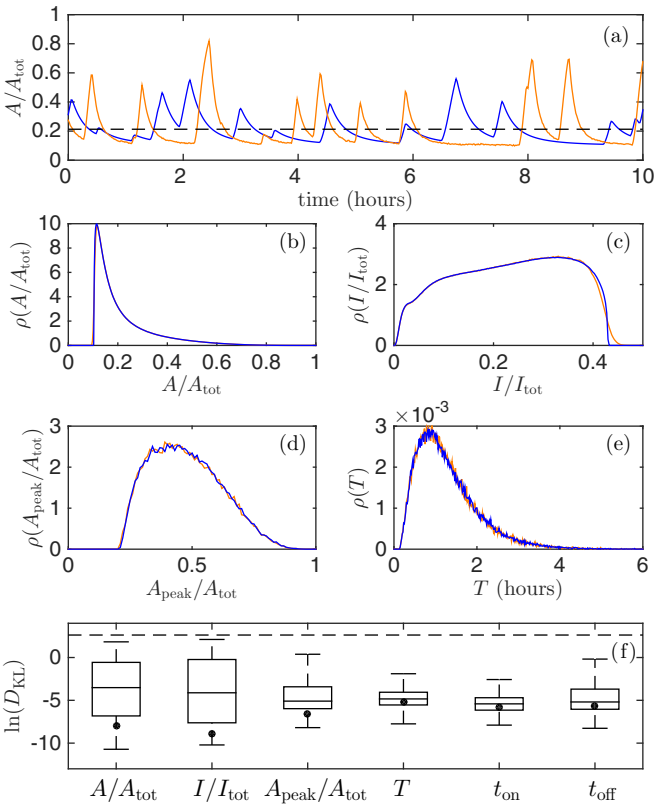

FIG. 2. (Color online) Stochastic gene switching is responsible for most of the dynamic variability in the network. (a) Evolution in time of the free activator A using fully stochastic [orange (light gray)] and simplified hybrid [blue (gray)] simulations. Both display similar pulses while the fully deterministic simulation (dashed line) remains at a stable fixed point. Distribution of the values of (b) A and (c) I for both the simplified hybrid [blue (gray)] and fully stochastic [orange (light gray)] simulations of the biochemical network. Distribution of (d) the amplitudes and (e) the periods of the pulses for the simplified hybrid [blue (gray)] and fully stochastic [orange (light gray)] models. The distribution of the activator and the pulse parameters are very similar for the simplified hybrid and the fully stochastic systems. The chosen parameter values are shown in the caption of Fig. 1. (f) Box plot showing the Kullback-Leibler divergence $D_{\rm KL}$ between the distributions of the magnitudes considered obtained from simplified hybrid simulations and fully stochastic simulations for 1000 combinations of random parameters; the box represents the 25th, 50th, and 75th percentiles while the lines above and below the box represent the 1st and 99th. Dots represent the $D_{\rm KL}$ values obtained for the parameter combination used in (a)–(e) and the 1000 additional parameter combinations were obtained by varying them randomly up to one order of magnitude above and below those values. The maximum $D_{\rm KL}$ value is displayed as a horizontal dashed line. Overall, for most parameter combinations $D_{\rm KL}$ values are low.

when the simplified hybrid distribution is used to approximate the stochastic distribution

$$D_{\rm KL} = \sum_i P_S(i) \ln \frac{P_S(i)}{P_H(i)}.$$

The results in Fig. 2(f) show that for the majority of the parameter combinations the $D_{\rm KL}$ values obtained are well below the maximum possible value [we found that it is approximately $\ln(1/\alpha)$, displayed in Fig. 2(f)] and for the parameters giving the results shown in Figs. 2(a)–2(d) the simplified hybrid simulations give a particularly good approximation to the stochastic simulations. Overall then, we have that simplified hybrid simulations faithfully reproduce the dynamics of the system for a wide parameter range.

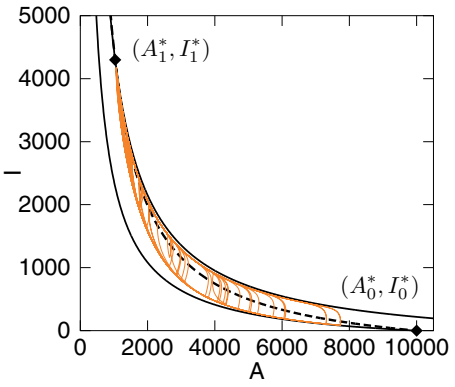

FIG. 3. (Color online) The three nullclines that can be associated with the simplified hybrid model: the nullcline for $\frac{dA}{dt} = 0$ (dashed line) and the nullclines for $\frac{dI}{dt} = 0$ with $G = 1$ (black line above) and $G = 0$ (black line below). A trajectory of the simplified hybrid model is also depicted [orange (light gray)]. There are two intersections between these three curves (black diamonds); a simple graphical analysis shows that these points are stable. Trajectories of the simplified hybrid model jump between these two fixed points. The chosen parameter values are shown in the caption of Fig. 1.

Although the dynamics of the network might depend strongly on the parameters considered, it is possible to make use of the point of view that our simplified hybrid simulation provides to show that the dynamics will follow qualitatively similar dynamics independently of the parameters considered. In particular, it is possible to understand the pulsed dynamics of our network in terms of the nullclines of the system given by Eqs. (7) and (8). Contrary to what we would have for a simple two-dimensional flow, here there are three of such nullclines: the one that we obtain by setting $\frac{dA}{dt} = 0$, $I = f_A(A)$, and the two nullclines that we obtain by setting $\frac{dI}{dt} = 0$ for $G = 0$ and 1, denoted by $I = f_{I,0}(A)$ and $I = f_{I,1}(A)$, respectively. The nullclines and a trajectory for our simplified hybrid model are depicted in Fig. 3. It is easy to see that, irrespectively of the parameter values, $I = f_A(A)$ intersects at exactly one point (A_0^*, I_0^*) with $I = f_{I,0}(A)$ and at exactly one point (A_1^*, I_1^*) with $I = f_{I,1}(A)$ (see Fig. 3). Furthermore, an analysis of the direction of the flow determined for each gene state shows that these two fixed points are necessarily stable. From this simple analysis we infer that the pulses can be understood as a series of jumps between the fixed points obtained for $G = 0$ and 1, as shown in Fig. 3, that will take place irrespectively of the parameters used in our network. However, as we will show in the following section, the parameters of the system have a far less trivial influence on the switching between these two fixed points, giving rise to different patterns of gene activation.

V. GENE ACTIVITY TIME DISTRIBUTIONS

In this section we focus on the patterns of gene activity arising from our simple model. Different experimental studies in which gene expression can be monitored in real time (either directly or by monitoring the expression of a short-lived protein or mRNA in time) have shown that gene active and inactive times can have different distributions. In particular, it has been observed that the gene active and inactive times can either be exponentially distributed or be described by peaked distributions [12–14]. These distributions are important signatures of the underlying stochastic process that drives gene activity. For instance, peaked distributions cannot be obtained from a simple stochastic switching process. To recapitulate the experimental observation, gene cycle models with multiple gene states have been proposed [12–14]. However, negative regulatory feedback loops could be an alternative mechanistic explanation to account for the observed gene dynamics.

To explore the gene active and inactive time distributions that our simple model can generate, we studied the dynamics of the system in parameter space by varying randomly each of the parameters one order of magnitude above and below the values used previously, given in the caption of Fig. 1. For each parameter combination, we simulated the dynamics and obtained the histograms of the active and inactive times $t_{\rm on}$ and $t_{\rm off}$ [see Fig. 1(b)]. We grouped them in ten clusters according to their coefficient of variation (CV), i.e., the standard deviation divided by the mean. We found that $t_{\rm on}$ and $t_{\rm off}$ can be distributed following quite distinct patterns: We observed distributions with shapes that range from an exponential-like shape, with a global maximum at zero and high CV, to a peaked Γ-like shape, with a global maximum close to the mean and low CV [see Figs. 4(a) and 4(b)]. Note that the Kullback-Leibler divergence between the gene activity distributions obtained from simplified hybrid and fully stochastic simulations is very low for the majority of the parameters combinations as shown in Fig. 2(f), indicating that our simplified hybrid approach approximates well the gene dynamics. Most importantly our simple system is able to recapitulate the experimentally observed distributions showing that negative feedback can give rise to peaked distributions and thus can be an alternative to the multistep process models proposed to explain the distributions observed in experiments [12,14].

To gain further insight into how the parameters shape the gene activity patterns, we examined the distributions of parameter values that produce exponential vs peaked distributions. Figure 4(c) shows that the parameters that discriminate best the two regimes are the association and dissociation rates k_a and k_d, the transcription rate k_I, and the gene switching rates $k_{\rm on}$ and $k_{\rm off}$. Interestingly, the peaked distributions emerged when the formation of the complex is favored ($k_{a,\rm exp} < k_{a,\rm peaked}$ and $k_{d,\rm exp} > k_{d,\rm peaked}$), in other words, when the negative feedback is stronger. On the other hand, the switching dynamics of the gene is slower ($k_{\rm on,exp} > k_{\rm on,peaked}$ and $k_{\rm off,exp} > k_{\rm off,peaked}$) for peaked distributions, indicating the importance of the interaction of the time scales at which the promoter and the negative feedback act in the gene activity patterns. Finally, peaked distributions emerge when the transcription of the inhibitor is more bursty, i.e., the number of molecules that are produced during the active periods is larger ($k_{I,\rm exp} < k_{I,\rm peaked}$) and the inactive periods are longer with respect to the degradation rate of the inhibitor ($k_{\rm on,exp} > k_{\rm on,peaked}$) making the transcription process more noisy and discontinuous.

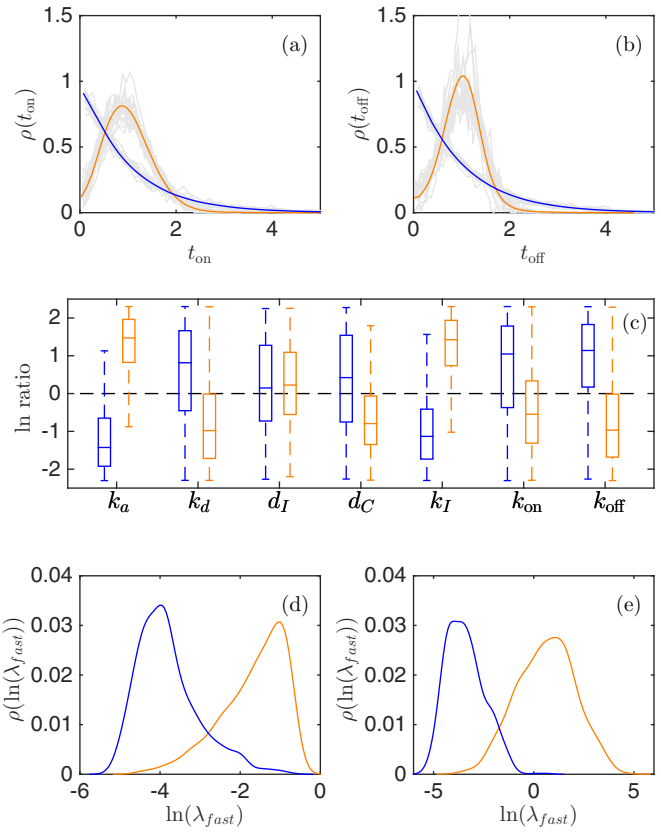

FIG. 4. (Color online) Distinct gene activity patterns are reproduced with different model parameters. Distributions of (a) the gene active times $t_{\rm on}$ and (b) the inactive times $t_{\rm off}$ with the highest and the lowest CVs [average in blue (gray) and orange (light gray), respectively] obtained by varying randomly the parameters given in the caption of Fig. 1 one order of magnitude above or below their value. (c) Box plot showing the distributions of the ratio between the parameter values and the reference values shown in the caption of Fig. 1, producing exponential distributions [blue (gray)] and peaked distributions [orange (light gray)]; the box represents the 25th, 50th, and 75th percentiles while the lines above and below the box represent the 1st and 99th. Distributions of the eigenvalues with larger absolute value $\lambda_{\rm fast}$ of (d) the fixed point (A_0^*, I_0^*) (obtained when the gene is inactive) and (e) the fixed point (A_1^*, I_1^*) (obtained when the gene is active), producing exponential distributions [blue (gray)] and peaked distributions [orange (light gray)].

Due to the high dimension of the parameter space, it is not feasible to predict the regions for which the different gene activity patterns are produced. However, our simplified hybrid approach also enable us to investigate the origin of these patterns using tools from dynamical systems. For the sake of simplicity we focus on the distribution of the active times $t_{\rm on}$. For our system, if at time $t = 0$ the system is in state G, the probability of remaining in state G can be expressed in terms of the conditional probability $P_G(t|V_0)$ given the initial condition $V_0 = (A_0, I_0)$ and the probability $P_{\bar{G}\to G}(V_0)$ of finding the

system at V_0 at the initial gene transition $\bar{G} \to G$ as

$$P_G(t) = \int dV_0 P_G(t|V_0) P_{\bar{G}\to G}(V_0), \tag{12}$$

where the conditional probability is

$$P_G(t|V_0) = k_{\text{off}} I(t,V_0) \exp\left(-\int_0^t k_{\text{off}} I(\tau,V_0) d\tau\right). \tag{13}$$

Notice that if the overall switching rate is constant ($\frac{dk_{\text{off}} I(t,V_0)}{dt} = 0$), as in the random switching model [11], an exponential probability distribution for the gene inactive time will be recovered. Instead, we find that in some cases the probability distribution is nonexponential. In this situation we would expect that the conditional distributions $P_G(t|V_0)$ that contribute the most to the integral in Eq. (12) should have a relative maximum at $t_{\max} \neq 0$. Using Eq. (13), we can see that such a relative maximum would satisfy

$$\left.\frac{dI(t,V_0)}{dt}\right|_{t_{\max}} = k_{\text{off}}[I(t_{\max},V_0)]^2. \tag{14}$$

Solving this equation requires the explicit form of $I(t,V_0)$. However, we know that $I(t,V_0) \to I_1^*$ and $\frac{dI(t,V_0)}{dt} \to 0$ as $t \to \infty$. Hence, Eq. (14) is the equation of the intersection of a monotonically decreasing function $\frac{dI(t,V_0)}{dt}$ [trajectories are always below the nullcline $I = f_{I,1}(A)$ and $\frac{dI(t,V_0)}{dt} > 0$; see Fig. 3] and a monotonically increasing function $k_{\text{off}}[I(t,V_0)]^2$ (because the amount of inhibitor grows when $G = 1$; see again Fig. 3) at the point $t = t_{\max}$.

It is possible to identify conditions allowing the fulfillment of Eq. (14) without knowing the explicit form $I(t,V_0)$. Our previous nullcline analysis shows that (A_1^*, I_1^*) is a stable fixed point with negative eigenvalues with absolute values $\lambda_{\text{fast}} > \lambda_{\text{slow}}$. Considering all this, we can roughly approximate

$$I(t,V_0) \approx I_1^* + (I_0 - I_1^*)(c_{\text{fast}} e^{-\lambda_{\text{fast}} t} + c_{\text{slow}} e^{-\lambda_{\text{slow}} t}),$$

where $c_{\text{fast}} > 0$, $c_{\text{slow}} > 0$, and $c_{\text{fast}} + c_{\text{slow}} = 1$. For short times (compared with $1/\lambda_{\text{slow}}$) the derivative $\frac{dI(t,V_0)}{dt}$ scales with λ_{fast}. Hence the crossing determined by Eq. (14), which is needed to have a peaked distribution, will only be possible if λ_{fast} is sufficiently big (and k_{off} is sufficiently small). It is easy to show that the same argument leads to an equivalent result for the distribution of the inactive times t_{off}, where λ_{fast} is the corresponding fast eigenvalue at the fixed point with $G = 0$.

We provide a numerical confirmation of the validity of this argument in Figs. 4(c) and 4(d), where we show that the values of λ_{fast} for the corresponding fixed points (for $G = 0$ and 1) are able to discriminate between the two different gene activity patterns, since the larger the eigenvalue is, the more likely that t_{off} and t_{on} are exponentially distributed. Note also that, according to Eq. (14) and our reasoning above, peaked distributions are more likely to occur for low k_{off} values, as confirmed numerically in Fig. 4(c). Thus, the perspective provided by our simplified hybrid simulations, by which all the reactions of the network can be approximated to evolve in a deterministic way while the gene switches stochastically, proves also to be a valuable tool to gain insight into the origin of the patterns of gene activity observed.

VI. CONCLUSION

It is becoming evident that pulsed dynamics is widespread in genetic circuits [34]. Our simple model shows that the interplay of negative feedback and stochastic gene switching gives rise to pulsed dynamics even if the fully deterministic simulations predict convergence to a steady state. In particular, the use of a simplified hybrid simulation scheme has allowed us to show that the stochasticity of gene activity has a drastic influence on the dynamics arising in this simple biochemical network and it can be responsible for most of the dynamical variability of the system. We believe that some of the ideas put forth in this paper based on simplified hybrid simulations, which allow one to combine notions of stochastic processes with those of dynamical systems, can help to gain insight into the dynamics of more complex genetic circuits. Furthermore, we have found that, in spite of the simplicity of the dynamics arising, our network can display different gene activity patterns. The negative feedback plays a key role in this fine temporal control: Without it, the dynamics of gene activation would be purely random. Our results imply that in experiments in which the gene activity patterns are found to be peaked [12–14] a negative feedback loop might be at work. We think that the increasing availability of experimental data will allow us to delineate the contribution to the gene activity dynamics by both the multistep sequential stochastic process of gene activation and the constraints imposed by the structure of the regulatory biochemical network. As in our work, the use of simplified hybrid simulations can help to provide further analytical insight.

ACKNOWLEDGMENTS

S.Z. was supported by the Intra-European Fellowships for career development through Grant No. 2011-298447NonLinKB. The work was supported by IEF Grant No. 2011-298447 (S.Z.) and Flagship Program Progetto Bandiera Epigenomica EPIGEN (MIUR-CNR) (M.E.B.). N.M. was supported by the Chancellor's Fellowships granted by the University of Edinburgh.

APPENDIX: HYBRID SIMULATIONS

To perform hybrid simulations of a system of biochemical reactions one essentially has to decide which processes can be approximated by the Langevin equation (and hence the number of copies can be well approximated by a continuous variable) and which have to be simulated as a discrete stochastic process (of species whose number can only be approximated by discrete variables) using the Gillespie algorithm [22]. The criteria can be summarized as follows [26]. Consider a system of N biochemical species with copy numbers $\{X_1, X_2, \ldots, X_N\}$ that interact through M biochemical reactions with probabilistic rates a_j, so the probability of the jth reaction taking place in dt is $a_j dt$. Let ν_{ji} be the change in species i due to the jth biochemical reaction. Then the reactions that can be numerically simulated as a Langevin equation with an integration step Δt should satisfy the following conditions [26]:

$$a_j(t)\Delta t > \alpha \gg 1 \tag{A1}$$

and

$$X_i(t) > \beta|\nu_{ji}|. \tag{A2}$$

The parameters α and β define how fine grained the variables have to be to appear as continuous valued; in the limit in which they tend to infinity, the approximation to the Langevin equation becomes exact. Condition (A1) means that approximation works well for reactions that take place many times in Δt, while condition (A2) means that the change in the reactants of the reaction considered should be relatively small if X_i has to be considered a continuous variable. Approximations to the Langevin equation work reasonably well for values of α and β close to 100 [26]. It is clear that for our simple genetic circuit and for any circuit involving a gene with just two possible states, the switching reactions can never satisfy Eq. (A2) for a β close to 100, so the approximation to the Langevin equation has to be excluded and gene switchings always have to be described by a master equation and numerically simulated using the Gillespie algorithm [25]. The remaining processes might then be modeled using Langevin equations. In the simplified hybrid approach that we exploit in this paper and has been used in a number of works [27–30] those processes are simply modeled using ordinary differential equations, ignoring then the fluctuations. We find that for our simple model this simplified hybrid approach provides a good approximation of the dynamics of the fully stochastic simulations obtained using the Gillespie algorithm [25] in a wide parameter range.

[1] U. Alon, *An Introduction to Systems Biology* (CRC, Boca Raton, 2007).
[2] A. Hoffmann, A. Levchencko, M. Scott, and D. Baltimore, Science **298**, 1241 (2002).
[3] N. Geva-Zatorsky, N. Rosenfeld, S. Itzkovitz, R. Milo, A. Sigal, E. Dekel, T. Yarnitsky, P. Pollack, Y. Liron, Z. Kam *et al.*, Mol. Syst. Biol. **2**, E1 (2006).
[4] D. Nelson, A. E. Ihekwaba, M. Elliott, J. Johnson, C. A. Gibney, B. Foreman, G. Nelson, V. See, C. Horton, D. Spiller *et al.*, Science **306**, 704 (2004).
[5] S. Zambrano, M. E. Bianchi, and A. Agresti, PLoS ONE **9**, e90104 (2014).
[6] G. Tiana, S. Krishna, S. Pigolotti, M. H. Jensen, and K. Sneppen, Phys. Biol. **4**, R1 (2007).
[7] F. Tostevin, W. de Ronde, and P. R. ten Wolde, Phys. Rev. Lett. **108**, 108104 (2012).
[8] J. E. Purvis, K. W. K. C. Mock, E. Batchelor, A. Loewer, and G. Lahav, Science **336**, 1440 (2012).
[9] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002).
[10] J. Peccoud, Theor. Popul. Biol. **48**, 222 (1995).
[11] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, PLoS Biol. **4**, e309 (2006).
[12] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, and U. Schibler, Science **332**, 472 (2011).
[13] C. V. Harper, B. Finkenstädt, D. J. Woodcock, S. Friedrichsen, S. Semprini, L. Ashall, D. G. Spiller, J. J. Mullins, D. A. Rand, J. R. E. Davis *et al.*, PLoS Biol. **9**, e1000607 (2011).
[14] N. Molina, D. M. Suter, R. Cannavo, B. Zoller, I. Gotic, and F. Naef, Proc. Natl. Acad. Sci. USA **110**, 20563 (2013).
[15] A. Coulon, C. C. Chow, R. H. Singer, and D. R. Larson, Nat. Rev. Gen. **14**, 572 (2013).
[16] L. Huang, Z. Yuan, P. Liu, and T. Zhou, BMC Syst. Biol. **9**, 1 (2015).
[17] A. R. Stinchcombe, C. S. Peskin, and D. Tranchina, Phys. Rev. E **85**, 061919 (2012).
[18] P. E. Morant, Q. Thommen, F. Lemaire, C. Vandermoere, B. Parent, and M. Lefranc, Phys. Rev. Lett. **102**, 068104 (2009).
[19] M. Assaf and B. Meerson, Phys. Rev. Lett. **100**, 058105 (2008).
[20] D. A. Potoyan and P. G. Wolynes, Proc. Natl. Acad. Sci. USA **111**, 2391 (2014).
[21] S. Zambrano, M. E. Bianchi, and A. Agresti, J. Theor. Biol. **347**, 44 (2014).
[22] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[23] D. T. Gillespie, J. Chem. Phys. **113**, 297 (2000).
[24] P. B. Warren, S. Tanase-Nicola, and P. R. ten Wolde, J. Chem. Phys. **125**, 144904 (2006).
[25] D. T. Gillespie, Am. J. Phys. **64**, 1246 (1996).
[26] H. Salis and Y. Kaznessis, J. Chem. Phys. **122**, 054103 (2005).
[27] S. Tay, J. J. Hughey, T. K. Lee, T. Lipniacki, S. R. Quake, and M. W. Covert, Nature (London) **466**, 267 (2010).
[28] P. Paszek, S. Ryan, L. Ashall, K. Sillitoe, C. V. Harper, D. Spiller, D. A. Rand, and M. White, Proc. Natl. Acad. Sci. USA **107**, 11644 (2010).
[29] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass, and M. W. Covert, Cell **150**, 389 (2012).
[30] T. Lipniacki, K. Puszynski, P. Paszek, A. Brasier, and M. Kimmel, BMC Bioinformatics **8**, 376 (2007).
[31] A. J. McKane and T. J. Newman, Phys. Rev. Lett. **94**, 218102 (2005).
[32] P. Rué and J. Garcia-Ojalvo, Math. Biosci. **231**, 90 (2011).
[33] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[34] J. H. Levine, Y. Lin, and M. B. Elowitz, Science **342**, 1193 (2013).